WHAT IS CLAIMED:

1. A method of comparing the semantic content of two or more documents, comprising:

accessing two or more documents;

performing a linguistic analysis on each document;

outputting a quantified representation of the semantic content of each document; and

comparing the quantified representations using a defined algorithm.

2. The method of claim 1, wherein the linguistic analysis comprises sentence analysis.

3. The method of claim 2, wherein the sentence analysis comprises a syntactic analysis and a semantic analysis.

4. The method of claim 1 wherein the quantified representation of a semantic content is a semantic vector.

5. The method of claim 4, wherein the semantic vector can have multiple components.

6. The method of claim 5, wherein each component can have multiple dimensions.

7. The method of claim 6, wherein each component of the semantic vector includes one or more text values.

8. The method of claim 7, wherein each text value can have one or more numerical values associated with it.

9. The method of claim 8, wherein each component of the semantic vector has three values:

a word or phrase appearing in the document or a synonym of said word or phrase;

a weighting factor associated with said word or phrase or synonym; and

a frequency value.

10. The method of claim 8 wherein each component of the semantic vector has two values:

a word or phrase appearing in the document or a synonym of said word or phrase; and

a weighting factor associated with that word or phrase.

11. The method of claim 4, wherein the semantic vector is a multi-dimensional vector defined by the content of a semantic net.

12. The method of claim 11, wherein the content of the semantic net is augmented by relative weights, strengths, or frequencies of occurrence of the features within the semantic net.

13. The method of claim 1, wherein the output of said defined algorithm is a measure of at least one of semantic distance, semantic similarity, semantic dissimilarity, degree of patentable novelty and degree of anticipation.

14. A method of comparing two or more documents, comprising:

linguistically analyzing two or more documents;

generating a semantic vector associated with each document; and

comparing the semantic vectors using a defined metric.

15. The method of claim 14, wherein said defined metric is one of:

$$\frac{\text{Sqrt}(f1^2 + f2^2 + f3^2 + f4^2 + \ + f(N-1)^2\ fN^2)}{n} * 100,$$

wherein f is a difference in frequency of a common term between two documents and n is the number of terms those documents have in common; or

Sqrt(sum((w-Delta)^2 * w-Avg)) / (Log(n)^3 * 1000),

wherein w-Delta is the difference in weight between two common terms, w-Avg is the average weight between two common terms, and n is the number of common terms, between two documents.

16. The method of claim 15, wherein a common term between two documents includes two terms that are synonyms.

17. The method of claim 14, wherein one or more of said two or more documents are located using an autonomous software or 'bot program.

18. The method of claim 17, wherein the 'bot program:

automatically analyzes each document in a defined domain or network by executing a series of rules and assigning an overall score to the document.

19. The method of claim 18, wherein all documents with a score above a defined threshold are linguistically analyzed.

20. The method of claim 14, wherein the semantic vector is a quantification of the semantic content of each document.

21. The method of claim 14, wherein the semantic vector can have multiple components, and each component can have multiple dimensions.

22. The method of claim 14, wherein each component of the semantic vector has a word or phrase appearing in the document or a synonym of said word or phrase; and

at least one of a weighting factor associated with said word or phrase or synonym and a frequency value.

23. A system for comparing two or more documents, comprising:

a document inputter, arranged to access two or more documents;

a semantic analyzer, arranged to perform a linguistic analysis on each document;

a semantic quantifier, arranged to output a quantified representation of a semantic content of each document; and

a comparator, arranged to compare the quantified representations using a defined algorithm.

24. A system for comparing two or more documents, comprising:

a document inputter, arranged to access two or more documents;

a semantic analyzer, arranged to perform a linguistic analysis on each document;

a semantic vector generator, arranged to output a semantic vector associated with each document; and

a comparator, arranged to compare the semantic vectors using a defined metric.

25. The system of claim 24, wherein said defined metric is one of:

$$\frac{\text{Sqrt}(f1^2 + f2^2 + f3^2 + f4^2 + \ + f(N\text{-}1)^2\ fN^2)}{n} * 100,$$

wherein f is a difference in frequency of a common term between two documents and n is the number of terms those documents have in common; or

Sqrt(sum((w-Delta)^2 * w-Avg)) / (Log(n)^3 * 1000),

wherein w-Delta is the difference in weight between two common terms, w-Avg is the average weight between two common terms, and n is the number of common terms, between two documents.

26. A computer program product comprising a computer usable medium having computer readable program code means embodied therein, the computer readable program code means in said computer program product comprising means for causing a computer to:

access two or more documents;

perform a linguistic analysis on each document;

output a quantified representation of a semantic content of each document; and

compare the quantified representations using a defined algorithm.

27. A computer program product comprising a computer usable medium having computer readable program code means embodied therein, the computer readable program code means in said computer program product comprising means for causing a computer to:

linguistically analyzing two or more documents;

generating a semantic vector associated with each document; and

comparing the semantic vectors using a defined metric.

28. The computer program product of claim 27, wherein the computer readable program code means in said computer program product further comprises means for causing a computer to:

identify one or more of said two or more documents using an autonomous software or 'bot program.

29. The computer program product of claim 27, wherein said 'bot program automatically analyzes each document in a defined domain or network by executing a series of rules and assigning an overall score to the document.

30. The computer program product of claim 27, wherein the semantic vector is a quantification of the semantic content of each document.

31. The computer program product of claim 27, wherein the output of said defined metric is a measure of at least one of semantic distance, semantic similarity, semantic dissimilarity, degree of patentable novelty and degree of anticipation.

32. The computer program product of claim 27, wherein said defined metric is one of:

$$\frac{\mathrm{Sqrt}(f1^2 + f2^2 + f3^2 + f4^2 + \ + f(N-1)^2\, fN^2)}{n} * 100,$$

wherein f is a difference in frequency of a common term between two documents and n is the number of terms those documents have in common; or

Sqrt(sum((w-Delta)^2 * w-Avg)) / (Log(n)^3 * 1000),

wherein w-Delta is the difference in weight between two common terms, w-Avg is the average weight between two common terms, and n is the number of common terms, between two documents.